MeSA-AE Validity Research

Background

The MeSA-AE Assist Test is based on the Trail Making Test that was originally part of the Army Individual Test Battery (1994). It consists of two parts and combined its two tests measure the mental processing speed, attention and mental flexibility. Specifically, Test A has been found in research to assess these cognitive abilities by using a complex visual scanning task (Schear and Sato, 1989). In Test B the subject has to connect number and letter sequences in an alternating pattern as quickly as they can. This subtest specifically requires divided attention. Test B is generally interpreted as a measure of executive functioning because it depends significantly on a person's working memory, attention and mental flexibility skills (Mitrushina, et al. 2005; Lezak and Howieson et al., 2012). The score for each subtest is determined by the total time it takes the individual to complete it. Errors are immediately addressed by stopping test takers and requiring them to begin again at the last correct number or letter. Consequently, errors increase an individual's test score time.

Research studies clearly support that the MeSA-AE Assist test is sensitive in identifying cognitive impairments resulting from mild TBI and dementia. In addition, this test has been found useful in detecting attention and concentration deficits for both children and adults. While the test is generally viewed as a non-specific measure of overall brain dysfunction it has been found to be a useful measure of attention and executive functioning, as well as, psychometric processing speed (Mitrushina, 2005).

Test Construct Validity

The MeSA-AE Assist Test uses the same test format and procedures with only minor modifications to the instructions as the Trail Making Test. The positioning and arrangement of both the two practice tests and the two main tests were constructed to have the exact same layout as the Trail Making Test. By using the same format, the MeSA-AE Assist test inherently has face validity. Research using alternative test forms similar to the ones in the MeSA-AE Assist was also found to have strong reliability coefficients with the standard test layouts (Charter et al., 1987; Franzen et al. 1986) supporting their use.

Clinical Validity and Applications

The Trail Making test is widely used by many psychologists, because it has been found to be a valid measure of the effects and severity of traumatic brain injury (Armitage, 1946; Spreen and Benton, 1965). Lange and Iverson et al. (2005) in their research found that the test performance

times increased proportionately based on the severity of the injury. An individual's performance on both tests was also found to significantly predict their level of independent functioning (M.B. Acker and Davis, 1989). In addition, Test A has been found to be sensitive to identifying individuals with a post-concussion syndrome (PCS). While an individual's score on Tests A and B were not found to be sensitive in significantly differentiating non-impaired individuals from those diagnosed with PCS, their high test time scores were clinically useful in ruling in PCS. This finding means that individuals who do not have PCS are unlikely to have impaired test time scores on either Test A or B (Cicerone and Azulay, 2002).

Lezak and Howieson et al., (2012) report that Test B has been found in several studies to be reflective of impairments in the left prefrontal cortex region. Both Tests A and B have also been found to be useful in identifying the beginning symptoms of dementia (Greenlief et al., 1985). Storandt, Botwinick et al. (1984) completed a research study that found Test A, by itself, was significantly useful in differentiating control subjects from individuals diagnosed with dementia. It has been also been found that older populations who are impaired in more complex daily living skills will show significantly higher test scores on Test B than non-impaired older adults (Bell-McGinty et al., 2002). The test has also been found to identify impairments for individuals who have been diagnosed with various types of dementia, including Alzheimer's disease (Chen et al., 2000). However, different types of dementia disorders were not found to be differentiated by their test scores and almost all individuals with dementia had high test time scores (Barr et al., 1992).

There are a number of other factors and disorders that are likely to result in higher test scores for both tests. Research has identified that individuals with emotional problems such as depression or more severe psychiatric disorders, including schizophrenia, are likely to show significant impairments on this test (Gass and Daniel, 1990; Crockett, Tallman, et al., 1988). Test errors were found to occur more often for individuals who had frontal lesions when compared to control subjects or individuals with posterior lesions (Stuss, Bisschop, et al., 2001). The test has also been found to be sensitive to identifying cognitive impairments for individuals who chronically abuse substances (McCaffrey, Krahula, et al., 1989).

The Spanish language version of the Trail Making Test was found to be valid for use with Hispanic populations using existing normative data criterion for evaluation of the test scores (Arnold, B.R., Montgomery, G.T. et al., 1994). While the interpretation of other neuropsychological tests was identified by these researchers to be biased for other ethnic populations, their findings do not support that this is likely to be true for the MeSA-AE Assist test. Thus, both Swedish and Spanish versions of this test can be validly interpreted using the normative data provided.

In summary, there is a significant body of research that supports the validity of this test in assessing impairments in general cognitive functioning, attention and mental processing speed that are likely to impact an individual's performance in both work and school environments. In addition, lower test quotient scale scores were found to be significantly indicative of greater impairment in individual's functioning. This test has also been found to be useful in

differentiating numerous disorders that are known to result in impaired cognitive functioning (Strauss, Sherman and Spreen, 2006).

Demographic Effects

Strauss, Sherman and Spreen completed a review of the research pertaining to demographic and cultural differences for this test. Gender differences in numerous studies were not found to reveal significant sex-related differences for adults. Slower score times were identified for Chinese-English bi-lingual and African American populations, but the primary relative differences in scores were found to be mostly attributable to the individual's age and secondarily to their education level. In general, these demographic variables were found to have less effect on test scores for clinical populations. Consequently, the MeSA-AE Assist Test normative scoring system uses age based norms and automatically adjusts the normative data set based on the person's education level, if available.

The Spanish language version of the Trail Making Test was found to be valid for use with Hispanic populations using existing normative data criterion for evaluation of the test scores (Arnold, B.R., Montgomery, G.T. et al., 1994). While the interpretation of other neuropsychological tests was identified by these researchers to be biased for other ethnic populations, their findings do not support that this is likely to be true for the MeSA-AE Assist test. Thus, both Swedish and Spanish versions of this test can be validly interpreted using the normative data provided for the MeSA-AE Assist Test.

Test Reliability

A number of research studies have been completed that evaluated the test-retest reliability of the Trail Making Test. As noted above, the test consists of Test A and Test B and the reliability coefficients differ somewhat for these two tests. The reliability coefficients for Test A ranged from a low of .46 to a high of .94 for non-clinical populations (Strauss, Sherman and Spreen, 2006). However, the reliability coefficients in the research literature for Test A were not found to be as strong for clinical populations, particularly for those individuals who were more impaired (e.g. diagnosed with schizophrenia, HIV or Epilepsy). The test-retest reliability coefficients for Test B were typically higher and ranged from .44 to .95 for non-clinical populations. The Test B coefficients fell in the strong to high range for clinical populations; with a low of .62 and a high of .87. Based on seven reliability studies of normal populations reviewed by Strauss, Sherman and Spreen (2006) it was determined by this author that the mean reliability for Test A and Test B was .66 and .73, respectively. Given this extensive body of test-retest reliability research, it is evident that both Test A and Test B have sufficient reliability to support their use in clinical assessments.

Noticeable practice effects do exist when individuals are re-tested after a relatively short interval such as a few days or a week. In one study by Stuss et al. (1988) found that on average for tests spaced one week apart that Test A decreased from two to six seconds (9% to 19% improvement depending on the age group) and that Test B decreased from six to ten seconds (12% improvement for two different age groups, but no significant change for one age group). In this study, older adults typically had higher test scores to start, but only showed significant practice effects for Test A. In general, longer test interval times for numerous studies show little or only very modest change in test scores for healthy adults (Strauss, Sherman and Spreen, 2006). One study that completed testing after an average of nine months for neurologically stable adults showed a practice effect of only about one second for Test A and almost four seconds for Test B (Dikman et al., 1999). This study found high correlations for Test A and Test B that were .79 and .89, respectively. In summary, the overall research supports that the test is a reliable measure of the cognitive skills that it assesses and has minimal practice effects that are likely to fall within a 90% confidence interval range. Thus, for the most part on retesting most individuals will not achieve decreases in their test completion time that reflect a meaningful change of their scores which cannot be reasonably accounted for by measurement error (Strauss, Sherman and Spreen, 2006).

Using Alternative Test Forms

Alternative forms of Test A and Test B have been found to have high correlation coefficients supporting their use. Charter et al. (1987) reported reliability coefficients for an alternative test that reversed the stimulus order ranging from .89 to .95 for Test A and from .92 to .94 for Test B. This study was completed using both normal and mix sample diagnostic groups. Overall, these reliability scores are equivalent to the scores obtained from the reliability studies using the original Tests A and B.

The design for developing alternative forms that has good reliability involves keeping the circle spatial layout the same, but rotating it either horizontally, vertically or in both directions. In this way the distance relationship between the various individual test stimuli is preserved. Essentially, this means that alternative test forms will need to be created as “mirror-like images” of the original test in order to assure that they have face validity. The three alternative versions provided in the MeSA-AE Assist are labeled Alternative 1, 2 and 3 and they maintain the same original test layout but were flipped horizontally, vertically and both horizontally and vertically, respectively. Based on the existing body of research these alternative versions can be used reliably for retesting.

The value of using an alternative form is that it helps minimize practice effects when testing individuals for less than three months. Based on the research completed by McCaffrey et al. (1993) practice effects are best controlled for by the administration of the test a minimum of

three months apart. This research found that all practice effects found from three preceding test administrations dissipated after three months' time. In cases when it is necessary that the MeSA-AE Assist test be administered more frequently than three months one of the three alternative test versions can be used to minimize the known practice effects.

